

# 基于动态缓冲阈值的 Internet 区分服务\*

林闯

(国家信息中心 北京 100045)

E-mail: clin@mx.cei.gov.cn

**摘要** 该文提出了一种将 Internet 网络分组传输延时和丢失控制的区分服务要求合并在一起的综合方案。在这个方案中,分组丢失控制是基于动态缓冲阈值(dynamic buffer threshold,简称 DBT)方案。在 DBT 方案中,缓冲控制阈值基于分组丢失行为而动态地变化。在静态缓冲阈值(static buffer threshold,简称 SBT)方案中,分组丢失变化没有控制,丢失率变化大。对于缓冲控制阈值的确定,没有简明的数学表达方式。由于 SBT 不能随着传输突发特性的变化而变化,从而影响网络的系统性能。该文提出的 DBT 方案不但可以克服这些问题,而且可以确定和控制标志分组与非标志分组丢失率之间的比例关系。文章给出了所提出方案的随机 Petri 网性能模型、不同分组级别之间的丢失比例数学公式以及此方案重要性能的模拟和分析结果。

**关键词** Internet, 区分服务, 动态缓冲阈值, 随机 Petri 网。

中图分类号 TP393

在 Internet 网络中,信息的传输控制一般不依赖于网络状态,仅提供单一的尽量做好(best effort)服务已不能满足各种网络应用的需要。在当前综合服务的 Internet 网络中,应有多种服务质量(QoS)的要求<sup>[1]</sup>,包括质量保证的(guaranteed)服务、可控负载的(controlled-load)服务<sup>[2]</sup>、区分(differentiated)服务和尽量做好服务。

要达到质量保证的服务是比较困难的,它需要复杂的资源预留选道、接纳控制和调度机制,这也隐含着在网络和用户(应用)之间需要复杂的接口<sup>[3]</sup>。更重要的是,质量保证的服务具有面向连结的特性,而 IP 网络的发展仍不具有面向连结的特性;更广泛地说,质量保证的服务面临着可扩展问题,网络规模的增加将导致质量保证服务复杂性迅速增加。

另一种提供服务质量控制的思路的方法是区分服务<sup>[4,5]</sup>。区分服务方法在网络内部节点仅依赖简单的存储器管理方法和调度算法,不依赖于连接通道,具有可扩展性。区分服务方案的基本思路非常简单,它在网络边界节点给一些分组印以标志,例如,可使用 IP 优先权域或者 TOS 字节,标志的分组在网络内部节点可以接受优先服务。区分服务虽然在理论上不能保证服务质量,但在链路带宽容量不断增加的情况下可以提供满意的服务质量。

目前,解决区分服务问题基本上有两种方案:确信服务(assured service,简称 AS)方案<sup>[6]</sup>和奖赏服务(premium service,简称 PS)方案<sup>[7]</sup>。AS 方案注重分组丢失的管理与控制,PS 方案注重分组实时传输调度。在 AS 方案中,可以采用缓冲阈值方法对不同级别分组的丢失进行控制,阈值一般是常数,在网络运行期间并不改变。这种控制方案在本文中叫做静态缓冲阈值(static buffer threshold,简称 SBT)方案。本文提出一个动态缓冲阈值(dynamic buffer threshold,简称 DBT)方案,其阈值随着分组丢失的变化而动态地调整到它的最佳值。网络的长时间运行性能与初始阈值的设置无关,DBT 方案可提供固定不同分组级别之间丢失比例的简单设置方法。SBT 方案的一个问题是分组丢失率的变化很大,没有得到有效的控制。在网络中,不但平均分组丢失率对信息传输质量

\* 本文研究得到国家自然科学基金(No. 69873012)、国家 863 高科技项目基金(Nos. 863-306-ZT05-01-02, 863-300-02-04-99)和国家重点基础研究发展规划项目基金(No. G1999032707)资助。作者林闯,1948 年生,博士,教授,博士生导师,主要研究领域为系统性能评价,计算机网络,随机 Petri 网,逻辑推理模型。

本文通讯联系人:林闯,北京 100084,清华大学计算机科学与技术系

本文 1999-08 09 收到原稿,2000-02-29 收到修改稿

是重要的,而且分组丢失率的变化也严重影响着信息传输的质量.在 DBT 方案中,将对分组丢失率的变化引入负反馈控制机制,使得分组丢失率的变化趋于平稳.另外,在 DBT 方案中,由于阈值随分组丢失的变化而调整,平均分组丢失率也会比 SBT 方案的相同情况有所改善.

服务质量问题主要包括信息传输的实时性要求和信息丢失的管理与控制等问题.因此,保证信息传输的实时性和丢失控制的综合要求是多媒体网络服务质量控制的一个重要研究问题<sup>[8,9]</sup>.在区分服务研究领域还缺乏这方面的工作,本文的一个重要目标是提供这方面的研究基础.

在 1997 年 12 月召开的 IETF 工作组会议中指出,区分服务下一步的一个重要问题是研究区分服务的性能模型和数值分析<sup>[10]</sup>.不同区分服务方案 and 不同实现机制的性能分析和比较对于区分服务方案的理解是关键的一步,它涉及到方案和控制参数的确定以及资源分配的管理原则.本文的研究工作的另一个目标是对区分服务的综合方案进行性能分析.我们将给出这种方案的随机 Petri 网模型,并提供这种模型的模拟分析和系统的性能参数.

## 1 建议方案的描述

本节将提出一个基于 AS 和 PS 方案的新的综合方案,此方案将吞吐量作为服务质量的一个特征,设置一个机制,直接控制用户期望的传输时间.

在网络边界节点,要解决信息分组的标志问题,它包括两个方面:如何实现在每一个应用中的分组标志以及如何区别不同应用之间的分组标志.每一应用分组标志的实现采用“漏斗”(leaky bucket)算法.

在网络内部的每一个节点都包括一个发送调度器和多个缓冲队列,如图 1 所示.不同类型实时传送要求的应用流使用不同的缓冲队列,而不是全体  $n$  类应用流共享,使用一个缓冲队列.在同一类型流中,对是否印有标志分组的缓冲管理采用部分缓冲共享(partial buffer sharing,简称 PBS)方案,在同一缓冲队列中,对是否印有标志分组的丢失控制设置不同的缓冲阈值和丢弃概率.任一缓冲队列  $i$  ( $1 \leq i \leq n$ ) 的最大空间为  $B_i$ ,队列从头向尾排序,是一个 FIFO 队列,如图 1 所示.

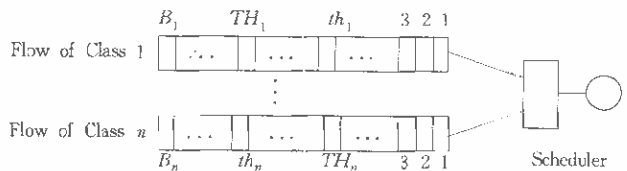


Fig. 1 Scheduling scheme model

图1 调度方案的描述模型

在模型中用两类变量参数来描述传输的调度和分组丢失控制.在流传输调度中,仅研究队列长度阈值(queue length threshold,简称QLT)调度方案,使用  $TH_i$  表示类型  $i$  ( $1 \leq i \leq n$ ) 流的实时调度缓冲阈值.当类型  $i$  流的分组在它自己的缓冲队列中的占有值达到  $TH_i$ ,并且更高实时优先级的类型流的队列缓冲占有值都小于它们自己的调度缓冲阈值时,则将发送权调度给类型  $i$  的分组缓冲队列.实时优先级越高的类型流,其调度缓冲阈值越小,反之亦然.

可以有如下两种分组丢失控制方案:

(1) 静态缓冲阈值(SBT).使用  $th_i$  表示类型  $i$  ( $1 \leq i \leq n$ ) 流的分组丢失优先级的缓冲阈值,其中  $1 \leq th_i \leq B_i$ .当非标志分组到达缓冲队列  $i$  时,如果缓冲队列的占有长度小于  $th_i$ ,则分组进入缓冲队列,否则,此到达分组被丢弃.到达的印有标志分组在缓冲队列没有满时进入缓冲队;仅当缓冲队列满时,到达的印有标志分组才被丢弃.换句话说,印有标志分组的丢弃缓冲阈值为  $B_i$ .

在这个方案中,分组丢弃不需要缓冲队列的扫描和到达分组的重新排序等复杂操作,因此,该方案实现简单.类型流缓冲队列可以根据用户要求动态地进行分配.

(2) 动态缓冲阈值(DBT).这个方案的实现比 SBT 复杂,对于类型  $i$  ( $1 \leq i \leq n$ ) 流的两个分组级别,各设置一个分组丢失计数器.计数器初值设置为 0.缓冲阈值  $th_i$  的初始值可设置为  $1 \sim B_i$  之间的任意值.每个计数器设置一个丢失增量,  $kl_i$  为级别 1(非标志分组)计数器  $CL_i$  所用的增量,  $kh_i$  为级别 2(标志分组)的计数器  $CH_i$  所用的增量.  $CL_i$  的每一个增量  $kl_i$  将导致缓冲阈值  $th_i$  增加一个值  $l_i$ ,亦即从  $th_i$  变成  $th_i + l_i$ .在另一方面,  $CH_i$  的每一

个增量  $kh_i$  将导致缓冲阈值减少一个值  $h_i$ , 亦即从  $th_i$  变成  $th_i - h_i$ . 这表明, 如果非标志分组丢失增加, 缓冲阈值就增加, 以达到增加缓冲的目的; 另一方面, 如果标志分组丢失增加, 缓冲阈值就减少, 以达到减少非标志分组缓冲的目的. 这个方案是一种缓冲阈值动态变化的方案, 缓冲阈值基于分组丢失行为而动态地变化.

### 2 方案的随机 Petri 网模型

假设读者对随机 Petri 网 (SPN) 的理论和应用有一些基本的了解. 有关 SPN 的描述可参阅文献 [11, 12].

#### 2.1 基于 SBT 方案的随机 Petri 网模型

在图 2 中, 我们给出了在网络内部节点基于 SBT 的传输实时调度方案的 SPN 模型, 此模型已经过精化设计<sup>[13]</sup>. 为了便于说明, 我们假定应用流类型实时传输的优先级按缓冲队列的次序反向排列, 亦即, 类型  $i$  流的实时传输优先级高于类型  $i+1$  流的优先级, 类型 1 流的实时传输优先级最高, 类型  $n$  流的实时传输优先级最低. 在该模型中, 字母下标的第 1 位表示实时传输的优先级, 第 2 位表示分组丢弃的优先级, 仅有一位下标表示实时传输的优先级. 为了简洁, 在模型中仅描述了类型 1 流和类型  $n$  流的子模型, 其他子模型相同. 在每一个类型流中, 分组丢弃优先级为两级.

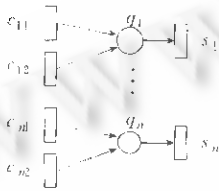


Fig. 2 A SPN model of the scheduling scheme based on SBT  
图2 基于 SBT 调度方案的 SPN 模型

在如图 2 所示的模型中, 有关变迁和位置的描述含意如下:

$c_{ij}$ : 表示类型  $i$  流的丢弃优先级  $j(j=1, 2)$  分组的产生, 优先级 1 的分组是非标志分组, 优先级 2 的分组是标志分组, 它们的产生速率呈指数分布, 其值为  $\lambda_j$ . 在同一类型  $i$  流中, 具有优先级  $j$  分组的产生比例为  $\lambda_j / \sum \lambda_j$ .

$s_i$ : 表示类型  $i$  流缓冲队列中的分组的发送服务, 其服务速率为  $\mu_i$ .

$q_i$ : 作为类型  $i$  流到达分组的缓冲队列, 其容量为  $B_i$ .

在如图 2 所示的模型中, 与变迁相关联的谓词描述如下:

$s_i$  的谓词描述流实时调度方案, 在本文仅讨论了 QLT 调度方案. 在 QLT 方案中,  $s_i$  的谓词为

$$[(M(q_i) \geq TH_i) \wedge (\forall j, 1 \leq j < i, M(q_j) < TH_j)] \vee [\forall j, 1 \leq j < i, M(q_j) = 0].$$

对于实时传输优先级最高的类型 1 流,  $s_1$  的谓词则简化为

$$(M(q_1) \geq TH_1) \vee (\forall j, 1 < j \leq n, M(q_j) < TH_j).$$

$c_{i1}$  的谓词为  $M(q_i) < th_i, 1 \leq th_i \leq B_i$ .

$c_{i2}$  的谓词为  $M(q_i) < B_i$ .

#### 2.2 基于 DBT 方案的随机 Petri 网模型

DBT 方案的子模型比 SBT 方案复杂, 任一系统  $i$  的 SPN 模型描述如图 3 所示. 在图 3 中的变迁  $c_{ij}$  和  $s_i$  以及位置  $q_i$  的含义与图 2 所示模型相同, 变迁  $c_{ij}$  没有实施谓词, 变迁  $s_i$  也实施 QLT 方案, 它的谓词与图 2 所示模型相同.  $y_{i2}$  表示类型  $i$  标志分组,  $y_{i1}$  表示非标志分组, 其他的瞬时变迁和位置的含义可描述如下:

$L_i$ : 模拟非标志分组的丢失. 当  $q_i$  的队列长度大于或等于阈值  $th_i$  时, 一个非标志分组到达, 它可实施, 实施后一个到达的非标志分组被丢弃, 一个标记增加到非标志分组丢失计数器位置  $CL_i$  中.

$H_i$ : 模拟标志分组的丢失. 当队列  $q_i$  满时, 一个标志分组到达, 它可实施, 实施后一个到达的标志分组被丢弃, 一个标记增加到标志分组丢失计数器位置  $CH_i$  中.

$I_i$ : 模拟丢弃缓冲阈值的增加. 当位置  $CL_i$  的标识值大于或等于  $kl$  时, 它可实施, 实施后位置  $CL_i$  的标识值减少  $kl$ , 位置  $th_i$  的标识值增加  $l_i$ .

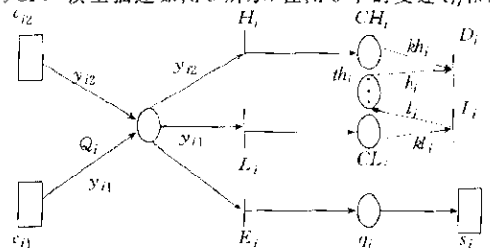


Fig. 3 A SPN sub-model of the DBT scheme  
图3 DBT方案的SPN子模型

$D_i$ :模拟丢弃缓冲阈值的减少.当位置  $CH_i$  的标识值大于或等于  $kh_i$  时,它可实施,实施后位置  $CH_i$  的标识值减少  $kh_i$ ,位置  $th_i$  的标识值减少  $h_i$ .

$E_i$ :模拟一个到达的分组被放入缓冲队列  $q_i$  中.

$Q_i$ :临时判断的保留空间位置.一个到达的分组首先进入位置  $Q_i$ ,以判定它是否被允许进入缓冲器(位置  $q_i$ ).

$th_i$ :非标志分组的丢弃缓冲阈值位置.这个位置的标识值表示当前的缓冲阈值.

$CL_i$ :非标志分组丢失计数器位置.

$CH_i$ :标志分组丢失计数器位置.

在图 3 中,瞬时变迁的实施条件谓词可以描述如下:

$L_i$  的谓词:  $M(q_i) \geq M(th_i)$ , 亦即  $q_i$  的占有量大于或等于缓冲阈值.

$H_i$  的谓词:  $M(q_i) = B_i$ , 亦即  $q_i$  满了.

$I_i$  的谓词:  $(M(CL_i) \geq kl_i) \wedge (M(th_i) + l_i \leq B_i)$ .

$D_i$  的谓词:  $(M(CH_i) \geq kh_i) \wedge (M(th_i) - h_i \geq 1)$ .

在图 3 中,变迁的实施优先级描述如下:

时间变迁  $c_{ij}$  和  $s_i$  有相同的优先级, 设置为 0.

瞬时变迁  $E_i$  有优先级 1.

瞬时变迁  $H_i, L_i, I_i$  和  $D_i$  有相同的优先级, 设置为 2.

阈值的移动频率是由  $kh_i$  和  $kl_i$  值决定的. 当  $kh_i$  和  $kl_i$  足够大时, 阈值不再变化, 那么 DBT 方案就退化为 SBT 方案. 因此, SBT 方案是 DBT 方案的特例, DBT 方案是更为广义的缓冲阈值方案. 阈值变化的大小是由  $h_i$  和  $l_i$  值决定的, 它们可以由应用流的突发性来决定. 突发性大的应用流, 可能需要更大的  $h_i$  和  $l_i$  值.

### 3 分组丢失比例和丢失变化

在基于 DBT 的方案中, 初始阈值的设置对于系统的稳态行为是不重要的, 阈值在系统运行期间可以动态地调整. 在不同分组级别之间的分组丢失率的比例是我们所关心的主要问题之一. 让  $P_{ij}$  表示类型  $i$  流的丢弃优先级为  $j$  分组的丢失率. 在模型中, 阈值在缓冲过程中朝高值位移动的速率是与  $kl_i$  和  $l_i$  值相关的. 变迁  $L_i$  的标记吞吐量为  $P_{i1} \times \lambda_{i1}$ . 在稳定状态下, 变迁  $I_i$  的标记吞吐量与  $L_i$  的吞吐量相同. 因此, 变迁  $I_i$  向位置  $th_i$  输送标记的吞吐量可表达为

$$P_{i1} \times \lambda_{i1} \times (l_i / kl_i). \quad (1)$$

同样地, 变迁  $D_i$  从位置  $th_i$  提取标记的吞吐量可表达为

$$P_{i2} \times \lambda_{i2} \times (h_i / kh_i). \quad (2)$$

在不考虑阈值在缓冲的两个端点(位置 1 和位置  $B_i$ )的运动受阻以及在相对稳定的状况下, 向左和向右运动的速率应相等. 因此可有以下方程:

$$P_{i1} \times \lambda_{i1} \times (l_i / kl_i) - P_{i2} \times \lambda_{i2} \times (h_i / kh_i). \quad (3)$$

当阈值达到缓冲的位置  $B_i$  时, 亦即非标志分组可以像标志分组一样使用全部缓冲时,  $P_{i1}$  和  $P_{i2}$  的比值可有最小值 1. 当阈值达到缓冲的位置 1 时, 亦即到达的非标志分组将全部被丢弃,  $P_{i1}$  和  $P_{i2}$  的比值可有最大值  $1/P_{i2}$ . 考虑到  $P_{i1}$  和  $P_{i2}$  的比值的最小和最大值, 方程(3)可以改写为

$$1 \leq P_{i1} \times \lambda_{i1} \times (l_i / kl_i) = P_{i2} \times \lambda_{i2} \times (h_i / kh_i) \leq 1/P_{i2}. \quad (4)$$

方程(4)允许我们设置参数  $kh_i, kl_i, h_i$  和  $l_i$  来控制不同级别分组之间的分组丢失率的比例. 进一步地, 方程(4)也告诉我们在 DBT 方案分组丢失率的变化是很小的. 当一个级别的分组丢失率变化时, 另一个级别的分组丢失率也将随之变化, 以维持预定的两个丢失率之间的比例值. 既然两个优先级别的分组共享一个缓冲器, 在两个级别的丢失率之间就存在着一个负反馈作用. 当一个级别的分组丢失率增加时, 阈值就将移动以增加它的缓冲空间, 减小它的丢失率. 这样就减少了另一个级别分组的缓冲空间, 也即增加了它的丢失率. 同样地, 另一个

级别的丢失率的增加,又能反过来控制第 1 个级别的丢失率,因此,在 DBT 方案中,两个级别分组之间的丢失率的比例趋向于稳定不变,而且分组丢失率的变化很小.在下面的数值例子中将显示 DBT 方案的这一特性.在实际系统中,两个级别之间的丢失比例是相当大的,一般要超过  $10^{-3}$ .因此,非标志分组的丢失率的变化相对于标志分组的丢失率变化来说,是比较小的.

这个负反馈作用的速率依赖于丢失率的比例和 4 个参数值  $kh_i, kl_i, h_i, l_i$  的设置.给定标志分组与非标志分组之间丢失率的比例和应用流输入的速率,我们可以设置这 4 个参数,使得丢失率比例满足要求,丢失变化尽可能小以及改善平均丢失率. $kh_i$  和  $kl_i$  比例的公倍数越小,负反馈作用的速率就越快.并不是在所有情况下,负反馈作用都是越快越好.在某些情况下,快速的负反馈作用可能造成平均分组丢失率的上升.因此,负反馈作用的快和慢要与应用流的传输特性相匹配.在我们的经验中, $kh_i$  值的范围为  $2 \sim 16, l_i$  为  $1 \sim 2$ .而且  $h_i$  一般设置为 1,  $kl_i$  由丢失比例决定.在上述范围内,可取得理想的结果.

#### 4 方案的性能评价和数值结果

在模型规模较小时,可以直接求解上述 SPN 模型.根据模型可以构造其对应的马尔可夫链.缓冲队列  $q_i$  的标识由变量  $b_i$  表示.模型的一个状态(标识)一般可由  $M[b_1, \dots, b_i, \dots, b_n]$  表示.状态  $M$  的稳定状态概率使用  $P[M]$  表示.

类型  $i$  流的丢弃优先级为  $j$  分组的丢失率  $P_{ij}$  可以表达为

$$P_{ij} = \sum P[M(q_i) \geq th(i, j)]. \tag{5}$$

其中,在基于 SBT 的方案中,  $th(i, 1) = th, th(i, 2) = B_i$ ; 在基于 DBT 的方案中,  $th(i, 1) = M(th_i), th(i, 2) = B_i$ .

类型  $i$  流的分组发送的吞吐量  $T(s_i)$  可以表达为

$$T(s_i) = \mu_i \times \sum_{M \in E} P[M], \tag{6}$$

其中  $E$  是使  $s_i$  可实施的所有可达标识的集合.

类型  $i$  流的分组在到达缓冲队列  $q_i$  时,如果发现有  $n$  个分组在队列中,那么系统的响应时间就是  $n+1$  个分组的的服务时间.类型  $i$  流的丢弃优先级为  $j$  分组在缓冲队列  $q_i$  的平均延时时间  $D_{ij}$  可以表达为

$$D_{ij} = \frac{1}{T(s_i)} \sum_{k=0}^{B_i-1} (k+1) \times P[M(q_i) = k] \times (1 - P_{ij}). \tag{7}$$

一般情况下,上述模型是一个多维的马尔可夫链,随着模型规模的扩大,系统的状态呈指数增长.由于此模型没有乘积形式的解,当状态数量超过一定限制后,当前一般计算机的存储和计算能力无法忍受时,可采用模拟方法求解.

由于要做大数量级的例子,在以下例子的计算和分析中,没有采用分析算法而采用了模拟计算方法.其重点在于比较 SBT 和 DBT 方案的平均分组丢失率和丢失变化方面的性能.我们给出两组例子:4 个参数的变化和影响以及分组丢失率的变化.这些例子都是基于如图 2 所示的 SBT 和如图 3 所示的 DBT 方案模型.

例 1:设置时间变迁的实施时间  $1/\lambda_2 = 0.2, 1/\lambda_1 = 0.1, 1/\mu_i = 0.06$ , 设置缓冲的容量  $B_i = 32$ . 阈值  $th_i = 26$  (对 DBT 来说,仅为初值).

SBT 方案模型的模拟结果为:标志分组的丢失率  $P_{i2} = 9.035834e - 06$ ; 非标志分组的丢失率  $P_{i1} = 8.777400e - 03$ ; 丢失的比例  $P_{i1}/P_{i2} = 9.713990e + 02$ .

在 DBT 方案模型中,可以根据 SBT 方案的丢失比例结果设置 4 个参数,获得相同的丢失比例,以便参照比较.表 1 给出了 4 组不同的 4 个参数设置所获得的结果.

Table 1 Results of parameter variations for DBT

表 1 DBT 方案模型不同参数设置的结果

	$kl_i$	$kh_i$	$l_i$	$h_i$	$P_{i2}(e-06)$	$P_{i1}(e-03)$	$P_{i1}/P_{i2}(e+02)$
1	1 940	1	1	1	8.196 771	8.064 665	9.838 832
2	3 880	2	1	1	7.772 261	7.633 078	9.820 924
3	3 880	2	2	2	8.893 820	8.749 474	9.837 700
4	7 760	2	2	1	7.707 845	7.593 796	9.852 036

由表 1 可知, 4 组不同参数的设置都按照方程(4)趋向同样的结果, 丢失比例与 SBT 情况相差大约为 1.5%。在 DBT 方案中, 所有平均分组丢失率都比 SBT 方案中的相应情况小 3%~15%。

例 2: 在这个例子中, 我们比较 SBT 和 DBT 方案的丢失率的变化。将整个模拟时间分成 15 个连续时间段(样板), 在每个时间段使用一个不同序列的随机数产生器来产生信源的传输, 但是模型的所有参数值保持不变。模型的参数设置为:  $1/\lambda_2 = 0.25, 1/\lambda_1 = 0.125, 1/\mu = 0.07143$ , 设置缓冲的容量  $B_i = 21$ , 阈值  $th_i = 14$  (对 DBT 来说, 仅为初值)。图 4 给出了标志分组丢失率变化在 SBT 和 DBT 方案之间的比较。可以清楚地看到, 在 SBT 方案模型中, 最大的分组丢失率和最小的分组丢失率之间相差超过 1 500 倍。但是在 DBT 方案模型中, 两者相差仅为 4%。非标志分组丢失率的变化如图 5 所示, SBT 方案和 DBT 方案的分组丢失变化都比较小。

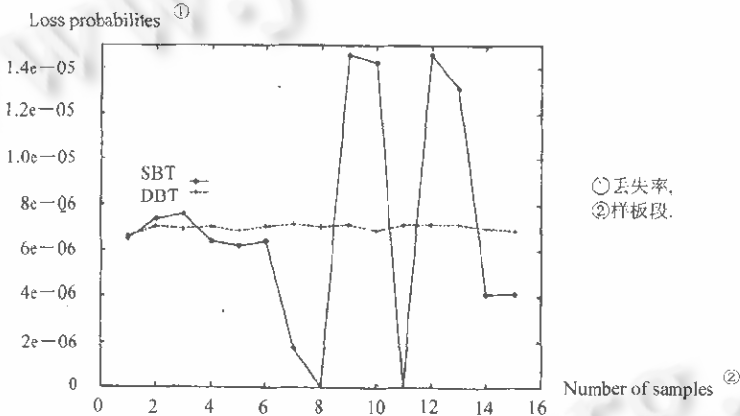


Fig. 4 Packet loss variation of marked packets

图 4 标志分组丢失率的变化

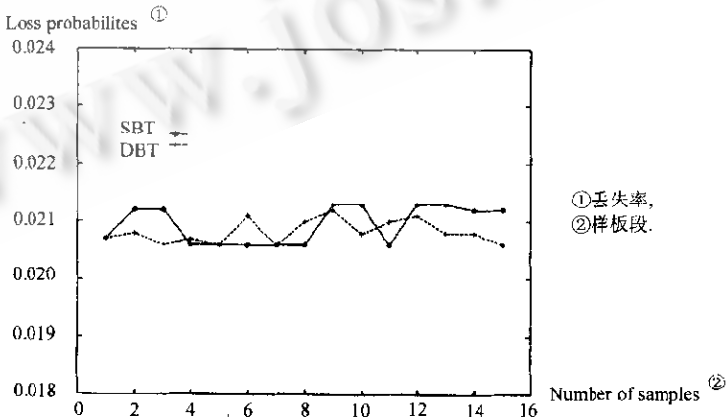


Fig. 5 Packet loss variation of unmarked packets

图 5 非标志分组丢失率的变化

### 5 结 论

本文描述了一种新的基于动态缓冲阈值 Internet 区分服务的实时调度和丢失控制相结合的方案,并给出了这种方案的 SPN 模型,简明的数学公式清晰地表达了不同分组级别之间的丢失比例和 4 个可控系统参数之间的关系.一些例子和经验结果表明, DBT 方案比 SBT 方案有更好的性能,特别是在分组丢失率的变化方面. 4 个可控参数为系统支持广泛的传输和不同的性能要求提供了一个丰富的工作环境. 本文所给出的区分服务性能模型和数值分析的结果对于区分服务综合方案的理解以及控制参数的确定和资源的分配具有指导意义.

### 参考文献

- 1 Ma Q, Steenkiste P. Supporting dynamic inter-class resource sharing: a multi-class QoS routing algorithm. In: IEEE INFOCOM'99. New York; IEEE Computer Society Press, 1999. 649~660
- 2 Wroclawski J. Specification of the controlled-load network element service. IETF RFC 2211, 1997
- 3 Lin C, Lam E C M. Dynamic queue length thresholds for scheduling real-time traffic in ATM networks. In: Proceedings of the 1999 International Conference on Communications. New York; IEEE Computer Society Press, 1999. 869~874
- 4 Blake S, Black D, Carlson M *et al.* An architecture for differentiated services. IETF RFC 2475, 1998
- 5 Nichols K, Blake S, Baker F *et al.* Definition of the differentiated services field (DS field) in the IPv4 and IPv6 headers. IETF RFC 2474, 1998
- 6 Clark D. Adding service discrimination to the Internet. Technical Report, MIT, 1995. <http://ana-www.lcs.mit.edu/anaweb/ps-papers/TPRC2-0.ps>
- 7 Nichols K, Jacobson V, Zhang L. A two-bit differentiated services architecture for the Internet. IETF Internet Draft (draft-nichols-diff-svc-arch-00.tex), 1997
- 8 Lin Chuang. Traffic scheduling based on the queue length threshold in ATM networks. Journal of Software, 1998,9(4): 316~320  
(林闯. ATM 网络基于队列长度阈值的传输调度. 软件学报, 1998,9(4): 316~320)
- 9 Lin Chuang. Integration of real-time scheduling and cell loss control in ATM networks. Chinese Journal of Computers, 1998,21(4): 333~340  
(林闯. ATM 网络一种实时传输调度和分组丢失控制的综合方案. 计算机学报, 1998,21(4): 333~340)
- 10 May M, Bolot J, Marie A *et al.* Simple performance models of differentiated services schemes for the Internet. In: IEEE INFOCOM'99. New York; IEEE Computer Society Press, 1999. 1385~1394
- 11 Lin C, Marinescu D C. Stochastic high-level Petri nets and applications. IEEE Transactions on Computers, 1988,37(7): 815~825
- 12 Ciaodo G, Muppala J, Trivedi K S. SPNP: stochastic Petri net package. In: Proceedings of the Petri nets and performance models. Los Alamitos, CA; IEEE Computer Society Press, 1989. 142~151
- 13 Lin Chuang. On refinement of model structure for stochastic Petri nets. Journal of Software, 2000,11(1): 104~109  
(林闯. 随机 Petri 网模型的精化设计. 软件学报, 2000,11(1): 104~109)

## Differentiated Services for the Internet Based on Dynamic Buffer Thresholds

LIN Chuang

(State Information Center Beijing 100045)

**Abstract** A new integrated scheme of differentiated services for the Internet is proposed in this paper. In this scheme, a service which combines both delay and drop priority is provided to applications with clear service differentiation. The integrated scheme is based on the dynamic buffer threshold (DBT) scheme. In the DBT scheme, the threshold is dynamically varied in run-time based on the packet loss behaviors. In the static buffer threshold (SBT) scheme, the variation of packet loss probabilities can be very large and there is no simple mathematical expression relating the threshold to the desired packet loss rate. These problems are overcome with the DBT scheme. In this paper, the performance model of the scheme is given using stochastic Petri nets and is analyzed by simulation.

**Key words** Internet, differentiated service, dynamic buffer threshold, stochastic Petri net.